图书情报工作
第63卷 第22期 2019年11月

基于在线评论的图书消费者满意度影响因素与作用机理*

尹丽春[1,2] 王悦[1]

[1] 黑龙江八一农垦大学经济管理学院 大庆 163319 [2] 山东青年政治学院政治与公共管理学院 济南 250000

摘要：［目的/意义］本文旨在提出一种从大量在线商品评论数据中挖掘影响读者满意度的关键因素的方法，并深入探讨各个影响因素对消费者满意度的影响模式和影响程度，进而为图书出版企业、电商平台持续改善读者满意度提供理论基础。［方法/过程］一方面利用朴素贝叶斯分类器将读者的情感进行分类。另一方面对评价文本中的高频名词进行聚类，发现影响读者满意度的主要因素。在此基础上基于最大程度减少不确定性的原则对各个影响因素的影响模式和影响程度进行分析。［结果/结论］以京东人工智能类图书评论为例进行了实证研究，发现包括内容、价格在内的六个因素可以极大地反映出读者的满意度(83.2%)。因此对于图书类商品，可以通过对大量历史评论数据的学习，找出影响读者满意度的主要因素，据此设计出简化的读者评论框架，以增强读者参与评论的积极性，提高评论的质量。“图书内容”是影响读者满意度的最主要因素。当读者对图书内容表达出不同的情感时，其他因素对读者满意度的影响模式和程度是完全不同的。当读者对图书内容表示满意时，89.2%的总体评论是好评，其他因素的影响较小，价格是导致中评和差评的最主要因素；当读者认为图书内容一般时，评论趋向于中性，读者对服务和物流更为关注；当读者认为图书内容不令人满意时，影响读者满意的因素依次是包装、服务质量和价格。基于不同影响因素对读者满意度的影响模式和影响程度，图书出版企业和电商平台可以更加有针对性地对其加以改善，以提高读者满意度。

关键词： 读者满意度　图书在线评论　机器学习　情感分析

分类号： G358　G252.1

DOI: 10.13266/j.issn.0252－3116.2019.22.012

1 引言

随着国民素质的提高，阅读书籍已经成为人们日常生活中必不可少的一部分，2018年亚马逊（中国）和“全民阅读”官网的共同调查显示年阅读量在10本以上的受访者，从2016年的48%增长到了56%。2017年全国图书零售市场总规模为803亿元，其中网络零售B2C市场图书出版物交易规模为301.3亿元，增速达到34.4%，如图1所示。预计2020年中国网络零售B2C市场图书出版物交易规模将达到646.1亿元[1]。

社交平台是读者分享阅读体验的主要渠道，基于社交网络的“一起读”方式成为数字阅读时代的独特行为。网络时代信息传播的速度之快，覆盖范围之广使得读者分享的阅读体验可以对大量的潜在读者产生影响。相关企业和电商平台纷纷通过积分或者小额代金券等形式鼓励消费者进行评论和分享。由此产生了大量的用户生成内容（User Generated Content，UGC），电商可以借此对客户进行精准画像，以实现精确营销的目的。同时还可以通过客户评价管理，吸引潜在消费者。

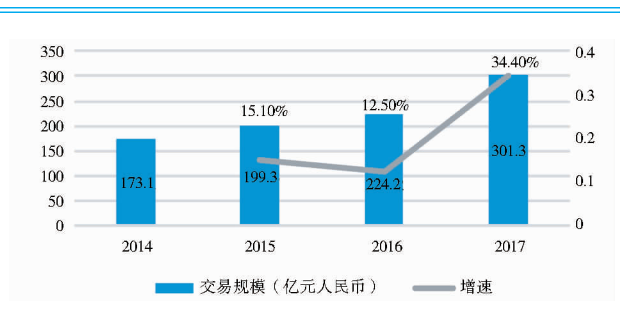


图1　中国B2C市场图书出版物交易规模[1]

* 本文系黑龙江农垦总局基金项目“大数据时代黑龙江垦区科技创新战略研究”（项目编号：HNK125B－14－08A）和黑龙江八一农垦大学研究生创新科研项目（项目编号：YJSCX2018－Y76）研究成果之一。

作者简介：尹丽春（ORCID：0000－0002－5510－1349），教授，博士，E-mail：308895957@qq.com；王悦（ORCID：0000－0003－4032－0263），硕士研究生。

收稿日期：2019－04－28　修回日期：2019－07－02　本文起止页码：106－117　本文责任编辑：杜杏叶

本研究以中国最大的 B2C 图书在线销售商城京东(JD. COM)(2017 年第 3 季度,京东 B2C 市场的图书在线销售以 36.2% 的市场份额超过当当和天猫,首次位居第一[2])的在线图书销售为平台,选取人工智能领域畅销书的在线评论为研究对象,对影响读者满意度的因素(评价的维度)进行判断,同时用机器学习的方法对评论的总体满意度进行分类(正向、中性和负向情感)。在此基础上对不同影响因素对读者满意度的影响模式和程度进行研究,以期为图书在线销售平台和出版企业进行图书在线评论管理提供经验证据,并为电商平台设计更加有效的评价系统提供支持。

2 理论基础

2.1 在线评论相关研究

在互联网技术盛行以前,人们常常会通过询问的方式来了解商品,这种行为被认为是线上评论发展的前身[3,4]。伴随着科技的迅猛发展,商品交易已经逐渐从线下转移到线上。与之对应,人们获取商品信息的渠道也发生了转移,由此产生线上评论。对于线上评论的概念,学者们普遍认同 C. Park 和 T. M. Lee[5] 的观点:消费者通过电商平台发表对所购商品的正负面评价。

经过多年的研究积累,在线评论的有用性得到广泛认可。M. L. Jensen 等提出:在线评论能够成为使用之前无法轻易表征的商品质量的可靠信息来源[6]。S. Ludwig 等人通过实验证明:比起商家的产品信息,网购消费者更加信任评论文本内容[7]。在线评论在塑造消费者对产品的认知和看法方面发挥着至关重要的作用[8-9]。

在线评论对产品销量的影响是研究的重点。学者们对多个领域的在线评论进行了研究。N. N. Hodac 等证实了亚马逊的在线商品评论会对产品的最终销量有一定程度的影响[10]。王君珺和闫强对京东手机评论的研究发现长文本评论对于各个品牌手机的销量都有直接影响[11]。J. H. Lee 等证实了在线评论与电影票房之间存在相关性[12]。W. Duan 和彭丽徽分别证实在线评论文本的数量与商品销售之间的正相关的关系[13,14]。消费者的购买决策行为对于企业的盈利能力也有重要影响[15],在线评论可以提高消费者对购物网站和产品的认知度,吸引潜在的消费者,从而使其做出更好的购物决策,以此提高消费者对企业产品的忠诚度[16]。

在线评论不仅能反映消费体验,在一定程度上还会影响到潜在消费者的购买意愿。在购买决策方面,某种产品或服务的在线评论显著影响着其他消费者的购买态度或行为[17],能够让消费者在做出购买决策之前对多种替代品进行评估和比较[18]。从消费者视角来看,在线评论能够为其在购买商品过程中提供所需信息,消除对商品本身或服务的部分顾虑[19]。

消费者的评论动机也取得一定进展。J. Heyman 与 D. Ariely 发现当购物网站对发表评论提供经济回报的情况下,评论者基于市场交换原则,更倾向于发布对购物网站或商家有利的评分[20]。J. C. Sweeney 等人在研究亚马逊网站图书消费者的评价动机时,发现经济回报会损害人们的内在动机,从而导致个人的努力程度下降,从而影响评价的有用性[21]。

由于消费者往往怀疑正向评论是消费者在获得经济回报的情况下,会与购物网站建立一种市场交换关系,从而发布对购物网站或商家相对有利的评论和评分[20],因此认为负面评价会提供更具有参考价值的线索,更有影响[22]。宁连举从程度、数量、内容相关性和专业性 4 个维度,研究了在线负面评论对网络消费者购买意愿的影响[23]。宋远征选取产品类型作为调节变量,将网络负面评论分解为无差异的产品属性负面评论、商家服务水平负面评论、物流负面评论,探讨其对购买意愿的影响[24]。

相比在其他传媒领域的应用,在线评论在图书出版业的应用还不够广泛。J. Chevalier 较早对 Amazon 和巴恩斯的在线图书销售数据进行研究,发现评分和数量都对结果产生积极影响,但评论长度对不同平台的影响显著性有所区别[25]。孙克琳以当当网不同种类图书评论数据为研究对象,证明了好评对图书销量有积极影响,差评对图书销量没有显著影响[26]。冀经纬利用回归分析对豆瓣网和当当网网络图书销量的影响因素进行了分析[27]。

2.2 消费者满意度相关研究

早在 20 世纪中叶,随着市场营销观念的盛行,一同备受关注的还有“消费者满意度”这一思想,但当时并没有学者进行系统的研究。到了 20 世纪 60 年代,美国学者 R. N. Cardozo 在探究消费者在购买商品前的投入成本和期望与购买商品后的满意程度间的关系时,正式提出“消费者满意度”这一概念[28]。并在实验中证明:当消费者花费相当多的精力来获得产品时,对产品的满意程度可能高于他们仅使用适当努力时的满意程度,而当产品达不到预期时,消费者满意程度低于产品达到预期时的满意程度。此后,对于消费者满意

度的探索就掀起了热潮,国内外众多学者纷纷从不同的视角来阐述消费者满意度的相关定义。

但到目前为止,学术界对消费者满意度没有统一的概念界定。R. L. Olive 在 1981 年和 1996 年分别提出了"消费者满意度是对先前期望和实际体验之间感知差异的总体情感反应",即预期不一致理论,和"消费者满意度是指消费者既定的心理需求得到充分满足并且商品及其附属服务令消费者获得愉悦情感的程度"理论,受到了学术界的普遍认可,这两种观点也是"消费者满意度"定义历史发展前后期的典型代表[29,30]。在 20 世纪 80 年代中期以前,学者们将消费者获得商品真实感受与期望的契合度作为主要研究方向。而在后期的探索中,有的学者认为消费者期望不是判断满意度的唯一指标,还有其他因素可以影响到消费者的满意度,如商品的本身和商品相关的服务等。消费者期望还存在由于对消费对象相关知识的匮乏,信息的缺失等,导致先前对商品期望过分夸张的现象。基于此,在 20 世纪 90 年代后,学者们研究的核心集中在满足消费者需求后的心理及情感表现方面。但无论是在前期还是后期,共同点都在于把消费者满意度理解为一种心理状态和情感表达。例如,H. K. Hunt 把消费者满意度定义为情感,而将消费者满意度定义为对情感的评价[31],H. Moezzi 则将消费者满意度总结为消费者对商品或服务体验评价后心理状态的结果[32],张圣亮等把消费者在消费过程中感知到的愉悦心理定义为消费者满意度[33]。

互联网的发展为网络购物开辟了巨大的空间,从根本上改变了企业与消费者之间关系的本质。因此消费者满意度的研究已经不仅仅局限于传统交易,网络购物消费者满意度也得到了国内外学者的重视。如,查金祥和王立生通过建立结构模型证实了线上商品的价格优势与质量对消费者满意度的影响最为显著[34];L. Zhang 等使用验证性因子分析和结构方程模型的方法,验证了计算机使用熟练程度,感知方便性和感知安全性是影响消费者对网上购物满意度的决定因素[35];李玉萍和胡培[36]以淘宝在线评论为研究对象,提出了商品品质、商家服务、快递服务质量等 8 个影响网络购物消费者满意度的因素;U. Tandon 等用北印度各州的 410 名在网购消费者的数据分析了网络购物环境中的消费者满意度的影响因素,揭示了网站功能、网站设计以及交易安全性等都对网络购物消费者满意度有一定的影响[37]。

与其他商品的消费者满意度研究相比,有关图书类商品的消费者满意度研究很少。R. T. Tapushe 发现信息和网站系统质量对在线新闻读者满意度的影响最大[38]。D. Saxon 等人从学生学习效果的角度,得出带超链接的期刊论文可以提高读者满意度[39]。张飞相在研究图书消费者忠诚度的时候,发现满意度对消费者忠诚度的影响最大[40]。匡霞在研究消费者感知质量与忠诚度关系时,发现消费者满意度在两者之间起到了中介的作用[41]。

基于以上文献研究分析可发现,因网络购物环境不同于传统的零售环境,所以消费者满意度的关键驱动因素也会不同,但其本质都是探究不同因素对消费者心理状态及情感的影响。而在网络购物环境中,在线评论作为消费者的指导信息,匿名在线评论的传播使更多的消费者愿意提供他们自己的真实体验与心理感受,表达对购买商品的情感倾向。

2.3 述评

在线评论作为大数据时代的重要数据源已经得到了较多的研究,其在商业上的价值已经获得了极大的认可。在线评论是消费者提供的最直观、最具体和最真实的消费体验,反映的是消费者最为关切的内容,因此是研究消费者满意度影响因素的重要资源。已有的研究中,在线评论情感分析、网购消费者满意度的影响因素都分别得到了较多研究,却没有从在线评论角度挖掘影响消费者满意度因素的研究。因此,本研究以京东商城中的人工智能领域的图书评论为例,从图书在线评论中挖掘影响读者满意度的因素。在此基础上进一步探索各个因素对读者满意度的影响模式和影响程度,为相关企业提高读者满意度提供理论基础和可操作的模式。

3 研究过程及结果

3.1 研究设计

本研究首先采用有监督的朴素贝叶斯分类器对在线评论的情感进行识别,然后采用无监督的 K-means 聚类算法寻找读者评论涉及的主要维度,作为满意度影响因素,最后利用决策树方法绘制出各个因素对读者满意度的作用机理,见图 2。

3.2 数据获取与预处理

3.2.1 数据获取 "人工智能"已经成为国际竞争的新焦点,不仅是引领未来的战略性技术,也成为经济社会发展的一个新引擎。京东图书平台 2018 年图书销量排行榜中的"计算机与互联网"目录下,排名前四名的都是与人工智能相关的书籍,它们分别是"Python 编

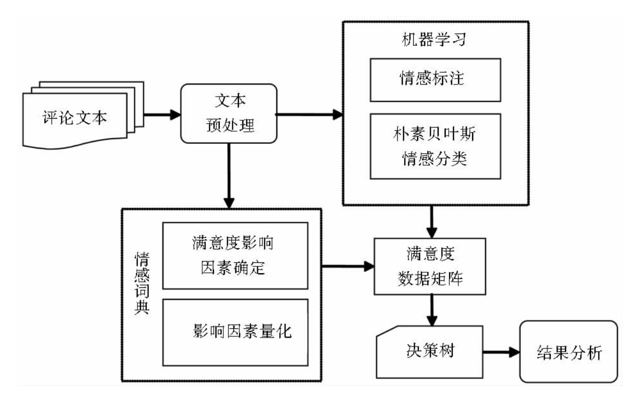


图 2　电商平台读者满意度影响因素研究框架

程－从入门到实践”“数学之美”“机器学习”和“深度学习”。本文以“人工智能”为关键词进行图书检索，发现了相关图书 64 294 本，其中有 123 本图书的评论数量超过 1 000 条。经过进一步地观察发现同一出版社的同类图书评论是被整合在一起的，因此评论数据搜集以出版社为单位进行，共获得 53 217 条评论文本。

3.2.2　数据预处理　众所周知，在线评论文本语句杂乱，结构性较差，因此在分析之前需要进行降噪处理。首先要对形式上的不当评论进行修改和剔除，例如繁体字转换为简体字、去重、去英文评论。经过降噪处理后还剩余 38 237 条评论，这些数据全部用于影响因素的提取。在进行机器学习时，考虑到单机的运行速度和学习的效果(不同情感的评论数量需要相对均衡)，基于多类情感均衡的原则从上述评论中抽取10 000条图书评论进行分析，经过人工甄别剔除无关评论后还剩余 9 729 条评论，以此为基础进行人工标注、情感分类和决策树的学习。

3.3　数据分析

3.3.1　评论文本整体极性分类　京东图书商城为了鼓励读者分享购物体验，对超过一定字数的评论给予“京豆”作为鼓励，虽然鼓励了大量消费者积极分享购物体验，但也存在少数消费者拼凑字数，提供无效评价的情况。虽然在数据处理过程中进行了初步处理，但仍然存在部分内容与评级不符的情况(见表 1)。这既是因为部分读者的评分标准不够准确，也反映出电商平台在设计读者评价体系和引导读者客观评价方面还有待改善。因此本研究首先对评论文本的真实情感进行分析。情感分析过程见图 3。

表 1　部分评级与评价内容不符的京东图书在线评论内容

id	购买时间	评论内容	读者打分
10727622675	2019－2－1 10:21	买来就跟旧书一样，有折痕；本地仓的结果又送到佛山再回到本地，不懂这个物流	5
10448981304	2018－10－23 15:11	不要买，真心挺一般的，书上的内容和别的书很多都雷同。基本内容网上也都能找到	5
10625154039	2018－11－29 9:08	这本书收到时很脏，申请了换货处理，依旧脏污，在等京东客服办理退货中，不打算要了	3
10517244172	2019－1－6 19:13	这本书内容中规中矩，是一本实用性很强的书，而且书的印刷质量也好，赞	3
10100666028	2018－12－20 12:17	颇为系统的一本书，覆盖了从机器学习到深度学习的完整知识体系，适合系统性学习或者高手随时翻查	1

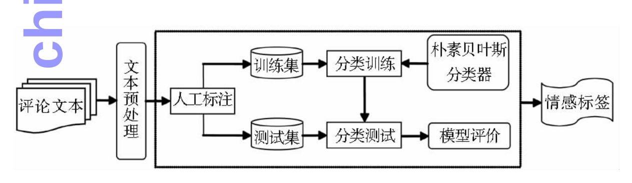


图 3　基于朴素贝叶斯的情感分类模型

(1)评论文本的极性标注。为了检验文本极性分类的准确性，本研究采用有监督的机器学习方法。首先对每一条评论的情感进行人工标注，将情感分为三级，分别赋值 1、2 和 3。赋值为 1 时包括以下几种情况：全部是负面的评论；正负评论都有，在不存在极端评价时，负面评论多于正面评论；存在极端评论时，极端负面评论多于极端正面评论。赋值 2 包括：全部是中性评论；正负面评论都有，在不存在极端评论时，正面评论多于或等于负面评论；存在极端评论，极端正面评论多于极端负面评论。赋值 3 包括：全部为正面评论；只包含对图书内容的描述。在这里，极端评价是指评价文本中出现了比较强烈的程度副词，如非常、及其、太等。

(2)文本结构化处理。评论文本属于典型的非结构化数据，因此在进行深入分析之前，需要将文本进行结构化处理。首先需要将文本以“词”为单位进行分割，然后构建出共词矩阵，最后再进行数字化转化，转变为词向量。

目前广泛应用的中文分词软件有中科院 ICTCLAS、JIEBA、清华大学的 THULAC 和 ANSJ 分词器等。其中 JIEBA 的应用最为广泛，分词正确率也较高。本研究借助 JIEBA 分词工具包进行结构化处理。

JIEBA 是一个通用型的分词软件，无法对不同的文本类型和文本主题进行准确分词，因此分词效果的好坏在很大程度上取决于针对文本特性和具体内容构建的自定义词典。读者发布的评论文本反映的是读者的购物过程体验，因此关注的焦点比较集中，表达的情

感相对明确。情感表达通常会存在着相对固定的搭配,而产品评论挖掘的主要任务就是挖掘评论中的情感词,因此本文借鉴陈硕总结出的表达情感的词性模板[42](见表2)对图书评论的情感词进行识别。在此基础上再进行人工分辨,进而形成针对图书评论的自定义词典,提高分词的效率和效果。

表2 图书评论中情感词性模板

词性模板	示例1	情感词	示例2	情感词
副词+形容词 da	非常不错	不错	有点脏	脏
名词+形容词 na	书不错	不错	内容丰富	丰富
副词+动词 dv	再买	买	没看	看
动词+形容词 va	送货快	快	感觉一般	一般
形容词+ 形容词 aa	好好	好	好薄	薄
助词+形容词 ua	就是快	快	应该不错	不错

基于表2情感词性模板提取出了图书评论文本中的情感词组合,对其中出现频次50次以上的情感词进行人工识别,进而形成了本研究的自定义词典。高频情感词如表3所示:

表3 图书评论中的主要情感词及出现频率

情感词组合	频率	情感词	频率
很好	802	还没看	207
还不错	307	不知道	165
不错不错	268	还没	250
都是	264	很好很好	133
挺好	248	质量不错	131
非常好	219	书不错	107

在分词的基础上可以构建图书评价的共词矩阵。由于不同的评论包含的词汇数量是不同的,因此处在不同维度的向量空间中,需要进行维度统一化处理,最终的维度就是所有评论中不同词汇的数量。

(3)情感极性分类算法选择。读者评论浩如烟海,要对所有文本都进行人工情感标注既要花费大量人力物力,也没有实际意义。本文采用可监督的机器学习算法对训练集的样本进行学习,再用训练好的模型对测试集的情感进行预测,根据预测值和标注值的一致情况检验算法的有效性。

朴素贝叶斯算法是应用最为广泛的分类算法,具有简单、高效的特点。本研究采用朴素贝叶斯算法对图书评论文本进行情感分类。朴素贝叶斯分类器(Naive Bayesian)是以统计学中的贝叶斯定理为基础(公式1)。根据贝叶斯定理,对一个分类问题,给定样本特征x,样本属于类别y的概率是:

$$p(y \mid x) = \frac{p(x \mid y)p(y)}{p(x)} \qquad 公式(1)$$

在这里,x是一个特征向量,假设x维度为M。因为假设特征条件独立,根据全概率公式展开,上述公式1可以表达为:

$$p(y = c_k \mid x) = \frac{\prod_{i=1}^{M} p(x^i \mid y = c_k)p(y = c_k)}{\sum_k p(y = c_k)\prod_{i=1}^{M} p(x^i \mid y = c_k)} \qquad 公式(2)$$

类别y的先验概率可以通过训练集算出。其中$P(y = c_k)$是类别c_k的先验概率。假设训练集中共有Q个样本,其中类别c_k包含$N(N<Q)$个样本,则有:

$$p(y = c_k) = \frac{N}{Q} \qquad 公式(3)$$

再通过训练集上的统计,分别估计出每一个特征在每一类的条件概率。$P(x^i \mid y = c_k)$是属性特征x^i在类别c_k中的类条件概率。指在类别c_k下,样本特征向量的某一维度下的某一特征出现的概率,假设c_k中有R个样本,在特征向量的第n个维度下特征x^i出现的次数为r,则有:

$$p(x^i \mid y = c_k) = \frac{r}{R} \qquad 公式(4)$$

由于分母中为全概率可视为常数,因此无需计算。

(4)分类效果评价指标。机器学习模型效果的好坏可以通过特定的指标来评价,其中最为经典的指标包括准确率(*Accuracy*)、精确率(*Precision*)、召回率(*Recall*)和*F1-score*。

$$Accuracy = \frac{TP\,真阳性 + TN\,真阴性}{TP\,真阳性 + FP\,假阳性 + TN\,真阴性 + FN\,假阴性} \qquad 公式(5)$$

$$Precision = \frac{TP\,真阳性}{TP\,真阳性 + FP\,假阳性} \qquad 公式(6)$$

$$Recall = \frac{TP\,真阳性}{TP\,真阳性 + FN\,假阴性} \qquad 公式(7)$$

$$F1_score = \frac{2 \times Precision \times Recall}{Precision + Recall} = \frac{2 \times TP\,真阳性}{2 \times TP\,真阳性 + FP\,假阳性 + FN\,假阴性} \qquad 公式(8)$$

其中,TP(True Positive)真阳性:预测为正,实际也为正;FP(False Positive)假阳性:预测为正,实际为负;FN(False Negative)假阴性:预测为负,实际为正;TN(True Negative)真阴性:预测为负,实际也为负。

Accuracy准确率是最常用的指标,反映的是整体的识别效果;Precision精确率体现的是"宁可漏掉,不可错杀"的思想,反映的是预测为正例的样本中有多少是正确的,数值越高,说明模型对正样本的识别能力越强;Recall召回率体现了"宁可错杀,不可漏掉"的宗

旨,也被称为“查全率”,反映的是真正的正例被识别出来的概率;F1-score是模型精确率和召回率的加权平均,它的最大值是1,最小值是0。F1-score越高,说明分类模型越稳健。

(5)分类结果比较分析。如果训练集中某个特征没有出现过,也就是一个词语在训练集中没有出现过,那么该词语的出现概率为0。使用连乘计算文本出现的概率时,会导致在整个文本中出现的概率也为0,这显然是不合理的。为了避免此种现象的发生,在估算概率值时通常要采用“拉普拉斯平滑($\alpha=1$)”或“Lidstone平滑($0<\alpha<1$)”进行修正。平滑系数是一个“超参数”,无法通过学习得到,因此需要人为设定,最终根据评价指标择优选择。

在Python的学习库中提供了两种比较有效的特征向量提取方法,一种是Count Vectorizer,它主要考虑每个词汇在当前训练文本中出现的频率;另一种是Tfidf Vectorizer,不仅考量某一词汇在当前训练文本中出现的频率,还关注这个词汇在其他训练文本中出现的次数。

停用词是指在文档中出现频繁但实际意义不大的词,例如“的”“在”“和”等。可以将停用词理解为过滤词,就是在分词过程中过滤掉的词。本文在评论文本分词之后,人工识别高频词中的无意义词,对JIEBA自带停用词库进行了扩展。但停用词是否会对分类结果具有正向作用需要通过实验来证实。

为了获得最佳分类效果,本研究采用交叉验证的方法,对不同平滑系数α(0.2,0.4,0.6,0.8,1.0)、特征抽取方式、是否利用停用词的各种组合分别进行学习,最后选择效果最好的一组。

最终确定当α取0.4,采用Count Vectorizer进行特征抽取,不用停用词的情况下学习效果最好,学习的准确率达到82.8%。其他指标如表4所示:

表4 不同情感分类效果的精确率、召回率和F1_score

满意度	Precision	Recall	F1_score
1	0.849 3	0.822 0	0.835 4
2	0.854 6	0.833 2	0.843 8
3	0.854 5	0.846 8	0.850 6
平均值	0.85	0.82	0.83

可以看出,三个分类的精确率、召回率和F1-score的值都大于82%,说明分类效果较好,好、中、差评的分类都比较准确。其中得分为3,即好评的分类召回率和F1-score值都较其他两类数值高,说明对“满意”评价的识别“漏掉”的比例最低,且分类最为稳定。中评的“错杀”率比较低,即预测为中评的评论中实际为中评的比例最高。

3.3.2 读者满意度影响因素确定及量化

(1)满意度影响因素确定。如何客观地在海量评论文本中找出影响读者满意度的主要因素是本研究的一个核心问题。传统的满意度影响因素都是以前人的研究成果为基础提出假设,并通过分析调查问卷对假设进行验证。然而在大数据时代,获取研究样本的方式更加便利,数据更加完整。可以借助大量的在线评论文本提取出读者的关注焦点,并以此作为满意度研究的评价指标[33,43-44]。文本中频繁出现的高频名词可作为评价文本的特征指标[45]用以体现读者对产品的关注点,也就是影响读者满意度的因素。

满意度影响因素确定的具体过程为:评论文本分词(全部评论),利用Word2Vec对分词后的文本进行训练,将每一个词语映射到词向量空间,再利用K-means算法对词向量空间中的高频名词进行聚类。

本部分的分词更加偏重于对名词的提取,因此除了构建自定义词典外,还需要多次对停用词库进行补充,以适应图书评论的特殊性,提高分词的准确性。在此基础上进行了词性标注。由于词汇数量巨大,所构建的词向量空间维数巨大且稀疏,因此需要降维处理。本研究采用Word2vec进行降维和词向量矩阵的构建,以用于聚类分析。

means算法是无监督学习的一个重要算法,它遵守竞争学习规则(winner-take-all rule),算法思想主要是在给定的数据集中随机选取K个质心,在质心周围形成簇组,经过多次的距离计算,最终使簇内距离尽可能小,簇间距离尽可能大。

主要高频名词的词云分布见图4,词云中的字体大小与词频(见表5)成正比。可以看出,“内容”是读者关注的核心焦点,说明影响京东读者满意度的最主要因素是图书的内容。同时可以看到质量、印刷、速度、价格等也是影响读者满意的主要因素。

为更为客观地挖掘影响读者满意度的关键因素,本文利用Word2Vec对评论文本进行训练,采用分布式表征,将训练文本的每个词语用分布式权重表示,并得到一个关于图书评论的词向量模型。然后将50个高频名词作为特征向量,构建了聚簇中心,选取K个随机质心,利用K-Means聚类算法进行多次迭代,将其分为K个类。



图 4　图书评价中的高频名词的词云分布

表 5　主要高频名词及其分布

高频名词	频数	高频名词	频数	高频名词	频数	高频名词	频数
内容	4 275	京东	2 907	价格	2 152	正品	1 869
纸质	4 171	包装	2 541	服务	2 084	物流	1 817
质量	3 305	帮助	2 366	纸张	1 921	外包装	1 800
印刷	2 942	速度	2 180	实惠	1 886	纸	1 781

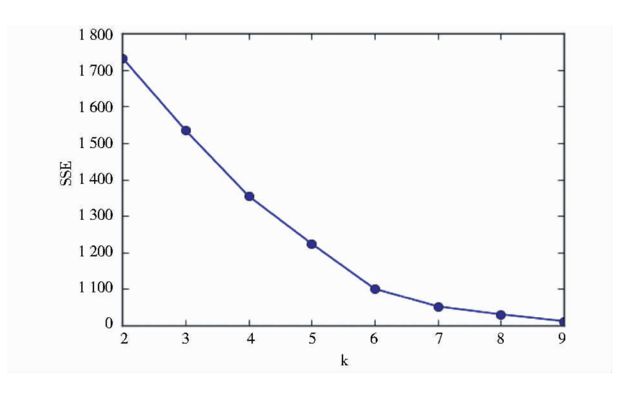


图 5　“肘部”法则

在 K 值选取上，通过观察可将 K 的取值范围锁定在 2 到 9 之间，利用肘部法则和轮廓系数同时验证聚类数 K 的值，多次运行后，得到最佳 K 值。

“肘部”位置即曲线拐点的位置对应的最佳分类数 K 值为 6，因此初步将 K 值确定为 6（见图 5）。接下来观察轮廓系数（Silhouette Coefficient）的运行结果，可以看到轮廓系数最多的 K 值是 6。根据轮廓系数的定义，最大值对应的 K 值为最优分类数（见图 6）。最后将空间向量逐步降维，得到图 7。通过层次聚类可将影响读者满意度的关键因素确定为 6 个。它们分别为：第三方服务相关的包装、物流、服务和与产品本身属性有关的价格、内容、和质量。

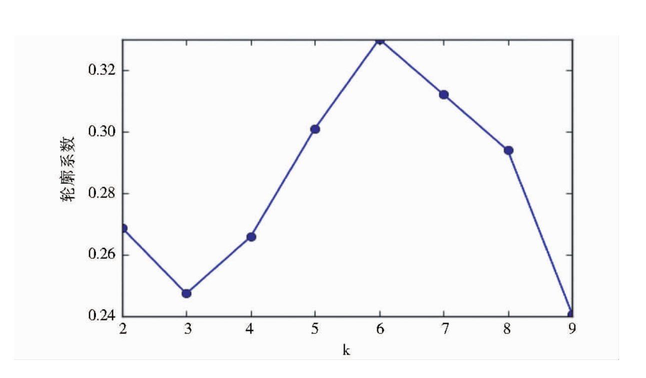


图 6　轮廓系数

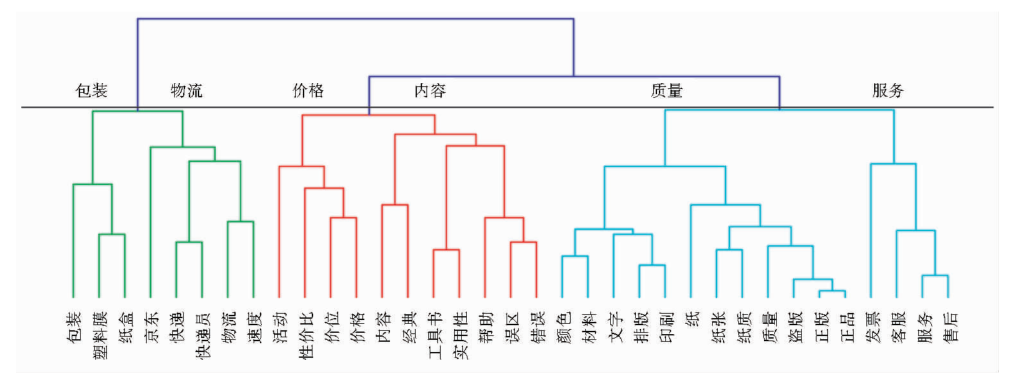


图 7　k = 6 时，高频名词分类效果

（2）满意度影响因素量化。用户对产品的评价与反馈，基本表现为对特征词的主观评价和描述。通过对大量的评论文本观察和分析发现，这种描述的结构可归纳为"特征词＋程度副词＋情感词"或"特征词＋情感词"。例如在评论"客服小姐姐人很好，服务态度超级棒，很详细地为我介绍哪本书适合我，而且书的印刷清晰，物流也很快，必须五星好评！"中就包含"服务态度＋超级＋棒"、"物流＋很＋快"和"印刷＋清晰"三种结构。因此评价指标量化就是以特征词为轴心，定位特征词前后［－u,u］的区间，然后对区间内的程度副词和情感词进行赋值，不同类型的文本 u 的取值是不同的，本文中 u 取值为 4。部分程度副词，情感词赋值如表 6 所示：

表 6　程度副词、情感词赋值（部分）

程度副词	分值	积极情感词	分值	消极情感词	分值	中性情感词	分值
非常、极其、太	4	好	2	差	－2	一般	1
很、挺	3	快	2	慢	－2	凑合	1
较	2	棒	2	劣	－2	还行	1
/（无修饰副词）	1	精彩	2	枯燥	－2	无感	1

根据赋值，计算特征词 i 的得分：

$$Score(i) = a * word(i) \qquad \text{公式(9)}$$

其中 a 为程度副词分值，$word(i)$ 为形容特征词 i 的情感词分值。

在情感分析研究中，有学者提出效价转换器的概念，即有部分词或短语的出现可能改变情感倾向[43]。常见的转化器为否定词转化器，如在特征词前后定义区间内存在否定词会直接影响文本整体语义，形成极性反转。例如："书的/内容/很/精彩"（正向），添加否定词后变为"书的/内容/不是/很/精彩"（负向）。因此，计算特征词 i 的得分：

$$Score(i) = (-1)t \ * a * word(i) \qquad \text{公式(10)}$$

其中 t 决定是否有极性反转。

根据以上读者满意度评价指标的量化方法，最终得到每条评论的每个评价指标特征词对应的分数。具体得分情况如表 7 所示：

表 7　评论文本特征词量化表（部分）

序号	评论内容	包装	物流	价格	内容	质量	服务
1	这本书写的真好，正版书，京东物流无敌，性价比高。	0	8	2	8	0	0
2	包装很差，就一个塑料袋。但好在书的内容不错，客服态度也比较好，勉强三星吧。	－6	0	0	2	0	4
3	这本书真的不建议购买，内容错误太多了，排版也不好。	0	0	0	－8	－2	0

最后，以每条完整在线评论为单位，汇总六类特征的情感分值，得到了满意度影响因素的多维数据矩阵。

3.3.3　基于决策树的读者满意度影响因素分析

（1）决策树算法选择。决策树方法就是一种从无次序、无规则的样本数据集中，根据决策问题的自然状态或条件出现的概率、行动方案、益损值、预测结果等，推理出决策树表示形式的分类规则的方法，用一个树状图表示出来，并利用该图反映出人们思考、预测、决策的全过程。

它采用自顶向下的递归方式。决策树从根节点开始，对实例的某一特征进行测试，根据测试结果将实例分配到其子节点，此时每个子节点对应着该特征的一个取值，如此递归的对实例进行测试并分配，直到到达叶节点，最后将实例分到叶节点的类中。因此从根节点到叶节点的一条路径就对应着一条规则，整棵决策树就对应着一组表达式规则[46]。

决策树学习通常包括特征选择、决策树的生成、决策树的修剪三个步骤。

决策树生成主要是由三个算法构成，分别是 ID3、C4.5 和 CART。3 个算法的主要区别在于度量信息方法、选择节点特征还有分支数量的不同。ID3 和 C4.5 采用熵（entropy）来度量信息不确定度，CART 采用基尼指数（Gini index）。ID3 选择"信息增益"最大的作为节点特征，C4.5 选择"信息增益比"最大的作为节点特征，它们都是多叉树。CART 是二叉树。

通过递归算法生成的决策树会根据已知数据一直计算下去，直到不能继续，很容易产生过拟合（overfitting）。剪枝算法就是提高决策树泛化性，防止过拟合的方法。剪枝分为预剪枝与后剪枝。预剪枝是指在决策树的生成过程中，对每个节点在划分前先进行评估，若当前的划分不能带来泛化性能的提升，则停止划分，并将当前节点标记为叶节点。后剪枝是指先根据训练集生成一颗完整的决策树，然后自底向上对非叶节点进行考察，若将该节点对应的子树替换为叶节点，能带来泛化性能的提升，则将该子树替换为叶节点。

（2）决策树学习结果及分析。本研究调用 Python

中的 Sklearn 数据库的决策树算法,采用优化 ID3 算法和预剪枝生成决策树,决策树的深度设定为 4。

随机选取满意度影响因素与对应的情感综合评价多维数据矩阵中的 75% 作为测试集,满意度影响因素作为 x-train,情感综合评价为 y-train。其余 25% 作为验证集,满意度影响因素为 x-test,情感综合评价为 y-test,预测结果为 y-pred。通过比对 y-test 和 y-pred,确定决策树的分类结果的准确率为 83.2%,分类效果比较好。其他指标如表 8 所示:

表 8 各影响因素与读者满意度关系的检验结果

	Precision	Recall	F1 – score
1	0.93	0.73	0.82
2	0.66	0.86	0.75
3	0.84	0.95	0.89
平均值	0.85	0.83	0.83

图 8 是全部样本的决策树。其中的每个内部节点代表对某一属性的一次测试,每条边代表一个测试结果。每个测试结果对应着“Ture”和“False”,左侧的为“Ture”,右侧的为“False”。决策树的决策过程需要从决策树的根节点开始,每一个内部节点都是以信息增益最大的影响因素作为判定依据,待测数据与决策树中的特征节点进行比较,并按照比较结果选择下一比较分支,直到叶子节点,作为最终的决策结果。

样本中,差评、中评和好评的数量分别为3 447、2 836和 3 446 条。在进行分析之前,整个系统的信息熵,也就是不确定性为 1.579 2。分别计算各个因素的信息增益,“内容”对读者满意度的影响最大,也就是说对内容的评价的认知会最大限度地降低读者满意度的不确定性。为了便于展示,基于对图书内容分值,将决策树分为两个部分(如图 8),上部分是对内容比较满意,评论数量为 2 161 条。下面部分是对内容满意度分值低于 1.5 的,评论数量为 7 568 条。

对于内容比较令人满意的图书(1.5 分以上),只要价格不是很贵(价格评分高于 –1.5)、服务水平一般及以上(服务评分大于 –0.5)、包装不是很差(包装评分大于 –1.5),89.2% 的评价得到好的综合评价(1 927/2 161 = 89.2%,路径①);虽然内容比较令人满意的评价中最终获得差评的样本数量极小,但也需要特别关注。可以看出在内容比较令人满意的前提下,价格是影响综合评价结果的一个重要因素。如果读者认为价格令人很不满意(得分低于 –3),无论其他因素如何,最终评价都没有好评②;对于其他因素,即使得分很低,读者会给与更多的中评,而不是差评,如过程③和过程④。

共有 1 964 条评论的内容得分介于 0.5 和 1.5 之间,即内容还行但还没有达到很好(⑤)。在这种情况下即使对服务、包装等影响因素存在不是很满意的情况,综合评价为中评的比例也很高,达到 88.7%(1 742/1 964 = 88.7%),但好评数量却低于差评。

内容一般及不好的评论数量最高,为 5 604 条。如果包装及图书质量没有达到很好及以上,58.8%(3 298/5 604 = 58.8%)的综合评价为差评。但对包装、质量和价格评分较高时,综合评价中也会有较高比例的好评(过程⑥及过程⑦)。

对于京东商城人工智能领域的读者来说,图书内容是影响读者综合体验最关键的因素。图书内容的感受与读者的综合评价具有明显的关联。说明人工智能领域的读者极其重视图书的内容,当图书内容一般或者较令人满意时,读者的综合评价基本与对图书内容的情感一致。但当内容不令人满意时,读者也比较宽容,只要在其他影响因素方面得到了较好的体验,就不会轻易给与差评。

不同的内容满意程度下,读者满意的决策过程也是不同的。对于内容很好的图书,价格太贵是导致差评的主要因素;服务和包装只要不是太差,对读者的满意度影响不大;图书质量和物流的影响更是微乎其微。内容还行、一般的图书,读者更加关注服务和物流,只要物流和服务不是太差,大部分的情感都是中评,但好评的数量低于差评数量。图书内容一般或者不好,只要包装和服务质量没有达到很高的水平,读者大部分都是不满意的;但包装、质量和价格的评分较高时,还是会有很大比例的读者会感到满意。

4 结论

本研究基于京东商城人工智能类图书的在线评论,提出了一种适应大数据环境的读者满意度的研究模式——结合机器学习和自然语言处理方法深入挖掘影响读者满意度的主要因素及这些因素如何影响读者满意度。文中所采用的研究方法具有较强拓展性,可以不受商品内容的限制,通过不断地学习自动提取特定商品的影响消费者满意度的影响因素,这些影响因素能最大限度地反映消费者对商品的总体评价。而且可以进行深度挖掘,通过在线评论分析,找到相关出版社在线图书销售中存在的问题。

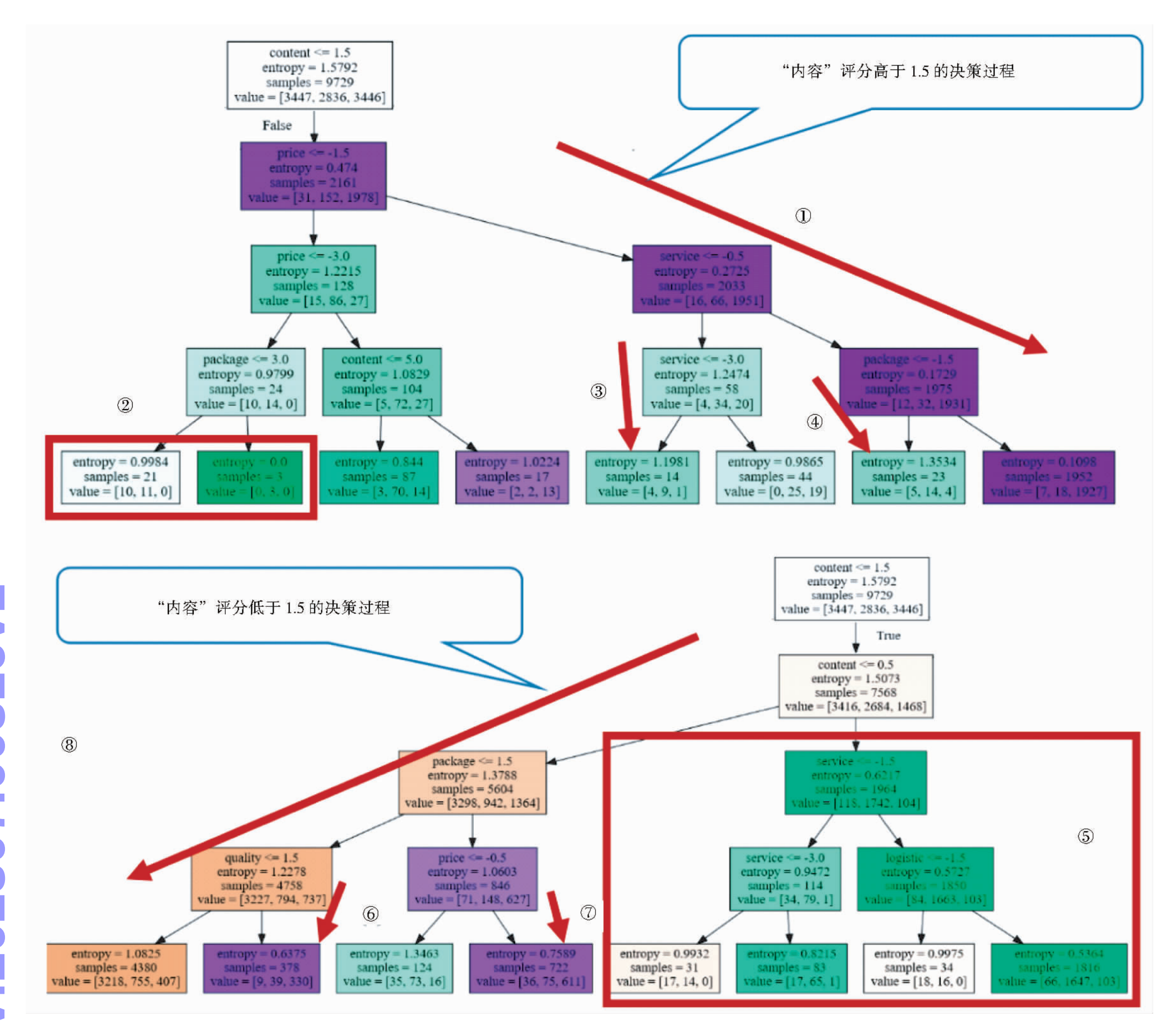


图 8　基于各个影响因素的读者满意度决策树

通过研究发现，朴素贝叶斯算法对评论文本情感的分类效果较好，学习的准确率达到 82.8%。这种方法能够解决读者在进行满意度评价时的非标准化和随意性问题，得到较为客观的综合评价分值。K-means 聚类算法也能有效地选取出反映读者满意度的主要维度。决策树能够直观地反映出各个因素对读者满意度的影响程度和影响模式。

对于京东商城人工智能领域的读者来说，包装、物流、服务、价格、内容、和质量是影响读者满意度的最关键的因素。不同的内容满意程度下，读者满意度的决策过程也是不同的。同理，电商平台也可以借助机器学习算法对其他在线商品的历史在线评论进行持续跟踪和分析，针对不同类型商品分别提取影响读者满意度的主要因素，从而简化评论流程，设计出更好的评价体系，引导读者评论。同时可以给出一个相对的标准，使读者有据可循，不仅降低了读者的评价难度，也提高了评价数据质量。基于消费者在不同情况下的差异的决策过程，商家和电商平台可以根据不同影响因素对消费者满意度影响模式和程度，有针对性地重点关注并及时改善，以提高消费者满意度。

参考文献：

[1] 当当，人人书. 2018 中国图书阅读市场专题分析报告[EB/OL]. [2018-06-10]. http://www.sohu.com/a/245466602_100218710.

[2] 易观. 中国 B2C 市场季度监测报告 2017 年第 3 季度[EB/OL]. [2018-06-10]. https://www.analysys.cn/article/analysis/detail/1000995.

[3] UTZ S, KERKHOF P, VAN D, et al. Consumers : how consumer reviews influence perceived trustworthiness of online stores [J]. Electronic commerce research and applications, 2012, 11(1): 49-58.

[4] TAX S S, BROWN S W, CHANDRASHEKARAN M. Customer evaluations of service complaint experiences: implications for relationship marketing [J]. Journal of marketing, 1998, 62(2):60-76.

[5] PARK C, LEE T M. Information direction, website reputation and eWOM effect: a moderating role of product type [J]. Journal of business research, 2009, 62(1):61-67.

[6] JENSEN M L, AVERBECK J M, ZHANG Z, et al. credibility of anonymous online product reviews: a language expectancy perspective [J]. Journal of management information systems, 2013, 30(1):293-324.
[7] LUDWIG S, RUYTER D, FRIEDMAN M, et al. More than words: the influence of affective content and linguistic style matches in online reviews on conversion rates [J]. Journal of marketing, 2013, 77(1): 87-103.
[8] LI X, HITT L M. Self-selection and information role of online product reviews [J]. Information systems research, 2008, 19(4): 456-474.
[9] XIAO S, WEI C P, DONG M. Crowd intelligence: analyzing online product reviews for preference measurement [J]. Information & management, 2016, 53(2):169-182.
[10] HODAC N N. CARSON S J, MOORE W L. The effects of positive and negative online customer reviews: do brand strength and category maturity matter? [J]. Journal of marketing, 2013,77(6):37-53.
[11] 王君珺,闫强.不同热度搜索型产品的在线评论对销量影响的实证研究[J].中国管理科学,2013,21(S2):406-411.
[12] LEE J H, JUNG S H, PARK J H. The role of entropy of review text sentiments on online WOM and movie box office sales[J]. Electronic commerce research and applications, 2017, 22:42-52.
[13] DUAN W, GU B, WHINSTON A B. Do online reviews matter? ——an empirical investigation of panel data [J]. Decision support systems, 2008, 45(4):1007-1016.
[14] 彭丽徽,李贺,张艳丰,等.基于品牌声誉感知差异的在线评论有用性影响因素实证研究[J].情报科学,2017,35(9):159-164.
[15] AHMED R R, SEEDANI S K, AHUJA M K, et al. Impact of celebrity endorsement on consumer buying behavior[J]. Ssrn electronic journal, 2015(16):12-20.
[16] KUMAR N, BENBASAT I. Research note: the influence of recommendations and consumer reviews on evaluations of websites [J]. Information systems research, 2006, 17(4):425-439.
[17] FLOH A, KOLLER M, ZAUNER A. Taking a deeper look at online reviews: the asymmetric effect of valence intensity on shopping behavior [J]. Journal of marketing management,2013,29(5-6): 646-670.
[18] ZHAO K, STYLIANOU A C, ZHENG Y. Sources and impacts of social influence from online anonymous user reviews [J]. Information & management, 2018,55(1):16-30.
[19] 汪旭晖,张其林,杜航.在线顾客评论对产品销量的影响:品牌强度和产品成熟度的调节作用[J].管理工程学报,2018,32(3):9-18.
[20] HEYMAN J, ARIELY D. Effort for payment a tale of two markets [J]. Psychological science,2004,15(11):787-793.
[21] SWEENEY J C, SOUTAR G N, MAZZAROL T. Word of mouth: measuring the power of individual messages [J]. European journal of marketing, 2012, 46(1/2):237-257.
[22] SKOWRONSKI J J, CARLSTON D E. Negativity and extremity biases in impression formation: a review of explanations [J]. Psychological bulletin,1989,105(1):131-142.
[23] 宁连举,孙韩. 在线负面评论对网络消费者购买意愿的影响[J]. 技术经济, 2014, 33(3):54-59,96.
[24] 宋远征. 网络负面评论对消费者购买意愿影响-产品类型调节作用[D].哈尔滨:哈尔滨工业大学,2016.
[25] CHEVALIER J, GOOLSBEE A. Measuring prices and price competition online: amazon. com and barnesandnoble. com [J]. Quantitative marketing & economics, 2003, 1(2):203-222.
[26] 孙克琳, 张心悦, 林萍夏, 等. 电子商务中在线评论对商品销售的影响——以当当网自营图书为例[J]. 电子商务, 2018(2):33-34.
[27] 冀经纬. 在线评论影响图书销量的实证研究[D].北京:中国地质大学,2018.
[28] CARDOZO R N. An experimental study of customer effort, expectation, and satisfaction [J]. Journal of marketing research, 1965, 2(3):244-249.
[29] OLIVER R L. Measurement and evaluation of satisfaction processes in retail settings [J]. Journal of retailing, 1981, 57(3):25-48.
[30] OLIVER R L. A Behavioral Perspective on the Consumer [J]. Asia pacific journal of Management, 1997, 2(2):285-286.
[31] HUNT H K. Consumer satisfaction: discussant comments [J]. Advances in consumer research, 1983, 10(4):262-262.
[32] MOEZZI H. Correlation between the determinants of customer relationship management system, electronic (e-CRM) customer satisfaction [J]. Research bulletin, 2009, 4(22):25-12.
[33] 张圣亮,李小东.网上购物顾客满意度影响因素研究[J]. 天津大学学报(社会科学版), 2013, 15(2):109-115.
[34] 查金祥, 王立生. 网络购物顾客满意度影响因素的实证研究[J]. 管理科学, 2006, 19(1):50-58.
[35] ZHANG L, TAN W, XU Y, et al. Dimensions of perceived risk and their influence on consumers' purchasing behavior in the overall process of B2C [J]. Lecture notes in electrical engineering, 2012, 111:1-10.
[36] 李玉萍, 胡培. 顾客网络购物满意度影响因素研究[J]. 商业研究, 2015, 57(1):160-165.
[37] TANDON U, KIRAN R, SAH A. Analyzing customer satisfaction: users perspective towards online shopping[J]. Nankai business review international, 2017,8(3):266-288.
[38] TAPUSHE R T, IMRAN M, MOHAMED E H, et al. To read or not to read: modeling online newspaper reading satisfaction and its impact on revisit intention and word-of-mouth interdisciplinary [J]. Journal of information, knowledge and management, 2018, 13(3):337-359.
[39] SAXON D, PEARSON A T, WU P. Hyperlink-embedded journal articles improve statistical knowledge and reader satisfaction[J]. Journal of graduate medical education, 2015, 7(4):654-657.

[40] 张飞相，杨扬，陈敬良. 图书在线消费者忠诚度影响因素的实证研究[J]. 中国流通经济，2014(11):87－93.
[41] 匡霞,杨扬. 图书在线消费者感知质量对忠诚度影响的实证研究[J]. 科技与出版,2019,4:151－156.
[42] 陈硕. 基于词性与句法的产品评论特征及情感词挖掘[D]. 河北:河北大学，2012.
[43] 吴维芳,高宝俊,杨海霞,等. 评论文本对酒店满意度的影响:基于情感分析的方法[J]. 数据分析与知识发现,2017 (03):62－71.
[44] 夏名首. 基于买家在线评价的购后评价特征数据挖掘[J]. 中国流通经济,2017(7):30－38.
[45] 李惠富，陆光. 多类型分类器融合的文本分类方法研究[J]. 计算机应用研究，2019，36(03):119－122.
[46] 韩忠明,段大高. 数据分析与 R[M]. 北京:北京邮电大学出版社,2014.

作者贡献说明:

王悦:提出研究思路,设计研究方案,数据的采集、清洗和分析,撰写并修改论文;

尹丽春:完善设计思路,分析数据,撰写、审阅并修改论文,论文最终版本修订。

Influencing Factors and Mechanism of Book Consumer Satisfaction Based on Online Comments

Yin Lichun[1,2] Wang Yue[1]

[1] School of Economics and Management, Heilongjiang Bayi Agricultural University, Daqing 163319

[2] School of Politics and Public Administration, Shandong Youth University of Political Science, Jinan 250000

Abstract: [**Purpose/significance**] This paper aims to propose a method to mine the key factors influencing reader satisfaction from a large amount of online commodity review data, and deeply explore the influence mode and degree of each influencing factor on consumer satisfaction, so as to provide a theoretical basis for book publishing enterprises and e-commerce platforms to continuously improve reader satisfaction. [**Method/process**] On the one hand, naive bayesian classifier was used to classify reader emotion; On the other hand, the high frequency nouns in the evaluation text were clustered to find the main factors that affect reader satisfaction. On this basis, the influence mode and degree of each influencing factor were analyzed based on the principle of minimizing uncertainty. [**Result/conclusion**] Taking the JD's artificial intelligence book review as an example, an empirical study was conducted. It was found that the six factors including content and price can greatly reflect the reader's satisfaction (83.2%). Therefore, for books and commodities, we can find out the main factors affecting readers' satisfaction by studying a large number of historical review data, and design a simplified reader review framework accordingly, so as to enhance the enthusiasm of readers to participate in the review and improve the quality of the review. 'Book content' is the most important factor affecting reader satisfaction. When readers express different emotions on the content of the book, the influence mode and degree of other factors on reader satisfaction are completely different. When readers are satisfied with the content of the book, 89.2% of the overall comments are favorable, while other factors have little influence. When readers think the book content is general, the review tends to be neutral, and readers pay more attention to the service and logistics; When readers think the content of books is not satisfactory, the factors affecting the satisfaction of readers are packaging, service quality and price in order. Based on the influence mode and influence degree of different influencing factors on reader satisfaction, book publishing companies and e-commerce platforms can be improved more specifically to improve reader satisfaction.

Keywords: reader satisfaction　book reviews online　machine learning　sentiment analysis